A Minimax Characterization for Eigenvalues of Hermitian Pencils. II

B. Najman*
Department of Mathematics
University of Zagreb
Zagreb, Bijenička 30, Croatia

and

Q. Ye
Department of Applied Mathematics
University of Manitoba
Winnipeg, Manitoba, Canada R3T 2N2

Submitted by Peter Lancaster

ABSTRACT

We establish a minimax characterization for extreme real eigenvalues of a general hermitian pencil $\lambda A - B$. The matrix A is allowed to be singular, so infinity may be an eigenvalue. It is also proved that the extremum can be taken over real subspaces if A and B are real.

1. INTRODUCTION

A matrix pencil $\lambda A - B$ is said to be regular if there is a nonsingular linear combination of A and B; to be nonsingular if A is nonsingular; to be hermitian if both A and B are hermitian; and to be real symmetric if both A and B are real and symmetric. Recently a minimax characterization has been established for some extreme real eigenvalues of nonsingular hermitian

*Supported by Fond za znanstveni rad Hrvatske.

655 Avenue of the Americas, New York, NY 10010 0024-3795/93/$6.00

pencils; see [2, 4, 5] (and [1] and [2] for infinite-dimensional extensions). In this paper, we continue the work in [5]; specifically, we shall prove

(1) that the minimax formulas for nonsingular hermitian pencils (Theorem 3.1 of [5]) are valid for general regular hermitian pencils;

(2) that the formulas are true with the subspaces in $\mathbf{R}^n$ for regular real symmetric pencils.

If a pencil is irregular, any number is an eigenvalue. Therefore such a pencil is not very interesting; we shall confine ourselves to the regular pencils. If a pencil is regular but singular (i.e., det $A = 0$), then besides some finite eigenvalues it has some infinite eigenvalues. In principle a regular pencil can be transformed into a nonsingular one, and then available results apply. However, it does cause a certain inconvenience to perform such a transformation, and the formulas may become more complicated. The purpose of (1) is to avoid doing such a transformation and demonstrate that the presence of infinite eigenvalues does not affect the minimax formulas. A real symmetric pencil can be regarded as a hermitian pencil, and then we have the minimax formulas, which use complex subspaces. It is always desirable in linear algebra to use the real field if the matrices are real.

Our second result shows that in the real case we can indeed take the extremum over real subspaces only. This is clearly a theoretical improvement and is also beneficial for numerical application.

In Section 2 we recall canonical forms for regular hermitian and real and symmetric pencils. We also introduce notation, following [5] as closely as possible. The results are also stated in Section 2. The proofs are given in Section 3.

2. STATEMENT OF THE RESULTS

We recall the fundamental results concerning the canonical forms of regular hermitian matrix pencils. For a positive integer m, $\varepsilon \in \{-1, 1\}$, and $\lambda \in \mathbf{C}$ we define the $m \times m$ matrices $J(\lambda, \varepsilon, m)$, $K(\varepsilon, m)$ and integers $\kappa_{\pm}(m, \varepsilon)$ as follows:

$$J(\lambda, \varepsilon, 1) = \varepsilon\lambda, \quad K(\varepsilon, 1) = \varepsilon \qquad \text{if} \quad m = 1,$$

$$J(\lambda, \varepsilon, m) = \varepsilon \begin{bmatrix} & & & & \lambda \\ & & & \cdot & 1 \\ & & \cdot & \cdot & \\ & \cdot & \cdot & & \\ \lambda & 1 & & & \end{bmatrix},$$

$$K(\varepsilon, m) = \varepsilon \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix} \quad \text{if} \quad m > 1,$$

and

$$\kappa_{\pm}(m, \varepsilon) = \begin{cases} \frac{1}{2}m & \text{if} \quad m \text{ is even,} \\ \frac{1}{2}(m \pm \varepsilon) & \text{if} \quad m \text{ is odd.} \end{cases}$$

For $a, b \in \mathbf{R}$, we define the $2m \times 2m$ matrix $\hat{J}(a, b, m)$:

$$\hat{J}(a, b, 1) = \begin{bmatrix} b & a \\ a & -b \end{bmatrix} \quad \text{if} \quad m = 1,$$

$$\hat{J}(a, b, m) = \begin{bmatrix} & & & & \hat{J}(a,b,1) \\ & & & \cdot^{\cdot^{\cdot}} & K(1,2) \\ & & \vdots & \cdot^{\cdot^{\cdot}} & \\ & \cdot^{\cdot^{\cdot}} & & & \\ \hat{J}(a,b,1) & K(1,2) & & & \end{bmatrix} \quad \text{if} \quad m > 1.$$

The following results are proved in [6] (see also [3] and [7]):

(a) Let $\lambda A - B$ be a regular hermitian pencil. Then there exists a nonsingular matrix X such that

$$\begin{aligned} X^*AX &= \operatorname{diag}\big[A_{-p}, \dots, A_{-1}, A_1, \dots, A_r, \\ &\qquad A_{r+1}, \dots, A_{r+s}, A_{r+s+1}, \dots, A_{r+s+t} \big], \\ X^*BX &= \operatorname{diag}\big[B_{-p}, \dots, B_{-1}, B_1, \dots, B_r, \\ &\qquad B_{r+1}, \dots, B_{r+s}, B_{r+s+1}, \dots, B_{r+s+t} \big], \end{aligned} \tag{1}$$

where $r = r_+ + r_-$, and for some $\varepsilon_j \in \{-1, 1\}$, $m_j \in \mathbf{N}$, $p, r, s, t \geqslant 0$,

$$A_j = J(0, \varepsilon_j, m_j), \quad B_j = K(\varepsilon_j, m_j), \qquad j = -1, \dots, -p,$$

$$A_j = K(1,1), \quad B_j = J(\lambda_j^+, 1, 1), \qquad j = 1, \dots, r_+,$$

$$A_j = K(-1,1), \quad B_j = J(\lambda_j^-, -1, 1), \qquad j = r_+ + 1, \dots, r,$$

$$A_j = K(\varepsilon_j, m_j), \quad B_j = J(\lambda_j, \varepsilon_j, m_j), \qquad m_j > 1,$$

$$j = r + 1, \dots, r + s,$$

$$A_j = K(1, 2m_j), \quad B_j = \begin{bmatrix} 0 & J(\lambda_j, l, m_j) \\ J(\overline{\lambda}_j, 1, m_j) & 0 \end{bmatrix},$$

$$\operatorname{Im} \lambda_j > 0 \qquad j = r + s + 1, \dots, r + s + t.$$

(b) If additionally A and B are real (hence symmetric), then we can find a real nonsingular matrix X such that (1) holds and that A_j, B_j are as in (a) except for $j = r + s + 1, \dots, r + s + t$, when

$$A_j = K(1, 2m_j), \quad B_j = \hat{J}(a_j, b_j, m_j), \quad b_j > 0.$$

Note that all the matrices in (1) are real in this case.

Clearly, the blocks $j = -1, \dots, -p$ correspond to the eigenvalue infinity, which has the signs $\varepsilon_{-1}, \dots, \varepsilon_{-p}$ and (possibly) Jordan chains associated with it. The blocks

$$j = 1, \dots, r_+$$

correspond to the (real) eigenvalues of positive type,

$$j = r_+ + 1, \dots, r$$

correspond to the (real) eigenvalues of negative type,

$$j = r + 1, \dots, r + s$$

correspond to the real nonsemisimple eigenvalues, and

$$j = r + s + 1, \dots, r + s + t$$

correspond to the nonreal eigenvalues.

In both cases (a) and (b) define

$$d^{\pm} = \sum_{j=1}^{p} \kappa_{\mp}(m_{-j}, \varepsilon_{-j}) + \sum_{j=r+1}^{r+s} \kappa_{\pm}(m_j, \varepsilon_j) + \sum_{j=r+s+1}^{r+s+t} m_j.$$

By the uniqueness of the canonical forms, d^+ and d^- are uniquely determined by A and B. It is easy to see that the definition of $d^{\pm}$ here is consistent with the one given in [5] for nonsingular hermitian pencils.

We remark that a regular real symmetric pencil has two canonical forms (1) according to cases (a) and (b), and hence two definitions of $d^{\pm}$ are produced. It is easy to see that, however, both canonical forms yield the same $d^{\pm}$.

Finally, we recall the following:

$$\sigma_{\min}^{+} = \sigma_{\min}^{+}(A, B) = \min\{\lambda_j : m_j > 2 \text{ or } \varepsilon_j = 1, \text{ for } j = r + 1, \dots, r + s\},$$

$$\sigma_{\min}^{-} = \sigma_{\min}^{-}(A, B) = \min\{\lambda_j : m_j > 2 \text{ or } \varepsilon_j = -1, \text{ for } j = r + 1, \dots, r + s\},$$

$$\sigma_{\max}^{+} = \sigma_{\max}^{+}(A, B) = \max\{\lambda_j : m_j > 2 \text{ or } \varepsilon_j = 1, \text{ for } j = r + 1, \dots, r + s\},$$

$$\sigma_{\max}^{-} = \sigma_{\max}^{-}(A, B) = \max\{\lambda_j : m_j > 2 \text{ or } \varepsilon_j = -1, \text{ for } j = r + 1, \dots, r + s\}.$$

We state our results in two theorems. We define the infimum (or supremum, respectively) over the empty set to be $-\infty$ (or $+\infty$).

THEOREM A. *Let $\lambda A - B$ be a regular hermitian matrix pencil, $\lambda_1^+ \leqslant \cdots \leqslant \lambda_{r_+}^+$ be the eigenvalues of positive type, and $\lambda_{r_-}^- \leqslant \cdots \leqslant \lambda_1^-$ be the* *eigenvalues of negative type.*

(a) *If* $\lambda_i^+ > \max\{\lambda_1^-, \sigma_{\max}^+\}$, *then*

$$\lambda_i^+ = \max_{\dim \mathscr{S} = n-(d^+ + i)+1} \inf_{\substack{x \in \mathscr{S} \\ x^*Ax > 0}} \frac{x^*Bx}{x^*Ax}.$$

(b) *If* $\lambda_i^+ < \min\{\lambda_{r_-}^-, \sigma_{\min}^-\}$, *then*

$$\lambda_i^+ = \min_{\dim \mathscr{S} = d^- + r^- + i} \sup_{\substack{x \in \mathscr{S} \\ x^*Ax > 0}} \frac{x^*Bx}{x^*Ax}.$$

(c) *If* $\lambda_i^- < \min\{\lambda_1^+, \sigma_{\min}^+\}$, *then*

$$\lambda_i^- = \min_{\dim \mathscr{S} = n-(d^- + i)+1} \sup_{\substack{x \in \mathscr{S} \\ x^*Ax < 0}} \frac{x^*Bx}{x^*Ax}.$$

(d) *If* $\lambda_i^- > \max\{\lambda_{r_+}^+, \sigma_{\max}^-\}$, *then*

$$\lambda_i^- = \max_{\dim \mathscr{S} = d^+ + r^+ + i} \inf_{\substack{x \in \mathscr{S} \\ x^*Ax < 0}} \frac{x^*Bx}{x^*Ax}.$$

This theorem has exactly the same form as Theorem 3.1 of [5]. The novelty here is that A can be singular, so that the pencil may have infinite eigenvalues. The presence of infinite eigenvalues causes a change in the index shift $d^{\pm}$.

Next we state the result for the real case. If A and B are real, the subspaces $\mathscr{S}$ in Theorem A are complex, which is inconvenient and unnatural. However, the above results hold also in real spaces:

THEOREM B. *If the pencil* $\lambda A - B$ *in Theorem A is real symmetric, then the formulas there hold with real subspaces* $\mathscr{S}$.

3. PROOFS

We prove parts (a) of Theorems A and B. The other parts can be proved either similarly or by applying (a) to a transformed pencil, as in [5]. Without loss of generality we assume that A and B are in their canonical forms [the right-hand side of (1)] and $\lambda_i^+ = 0$; if $\lambda_i^+ \neq 0$, we consider the pencil $\mu A - \tilde{B}$ with $\mu = \lambda - \lambda_i^+$, $\tilde{B} = B - \lambda_i^+ A$, noting that the numbers $p, r_\pm$, $s, t, m_j, \varepsilon_j, d^\pm$ are the same for the new pencil.

We start with three lemmas.

LEMMA 1. *Let* $A_0 = J(0, \varepsilon, m)$, $B_0 = K(\varepsilon, m)$, *and* $t > 0$. *Then there exists a* $\kappa_-(m, \varepsilon)$*-dimensional subspace* $\mathscr{T}_t$ *such that*

$$x^* A_0 x > t x^* B_0 x \quad \textit{and} \quad x^* B_0 x < 0 \tag{2}$$

whenever $x \in \mathscr{T}_t \setminus \{0\}$.

Moreover, $\mathscr{T}_t \cap \mathbf{R}^n$ *is a real* $\kappa_-(m, \varepsilon)$*-dimensional subspace of* $\mathbf{R}^n$.

Proof. First consider $m = 1$. If $\varepsilon = 1$, the statement is void. If $\varepsilon = -1$ and $\mathscr{T}_t = \operatorname{span}\{1\}$ ($= \mathbf{C}$), we find for $x \in \mathscr{T}_t \setminus \{0\}$

$$x^* B_0 x = -|x|^2 < 0, \qquad x^* A_0 x - t x^* B_0 x = t|x|^2 > 0.$$

If $m \geqslant 2$, we apply Lemma 3.3 of [5] to find a subspace $\mathscr{T}_t$ of dimension $\kappa_+(m, -\varepsilon) = \kappa_-(m, \varepsilon)$ such that for $x \in \mathscr{T}_t \setminus \{0\}$ we have

$$x^* K(-\varepsilon, m) x > 0 \quad \text{and} \quad x^*[J(0, -\varepsilon, m) - tK(-\varepsilon, m)]x < 0.$$

Noting that $K(-\varepsilon, m) = -K(\varepsilon, m)$, $J(0, -\varepsilon, m) = -J(0, \varepsilon, m)$, we conclude that $\mathscr{T}_t$ has the required properties. The last statement follows from the fact that the matrix W in Lemma 3.3 of [5] is real. ■

We denote the Rayleigh quotient by

$$r(x) = \frac{x^* B x}{x^* A x}$$

whenever the denominator is nonzero.

LEMMA 2. *Let* u *and* v *be vectors such that*

(i) $u^* A u > 0$,
(ii) $v^* B v < 0$,
(iii) $v^* B v < r(u) v^* A v$.

Then there exists $x \in span\{u, v\}$ *such that*

$$x^*Ax > 0 \quad and \quad r(x) \leqslant 0. \tag{3}$$

If A, B, u, *and* v *are real, then* x *can be chosen real.*

Proof. If $r(u) \leqslant 0$ then set $x = u$. It remains to consider $r(u) > 0$. Set $x_\alpha = u + \alpha v$, $b(\alpha) = x_\alpha^* B x_\alpha$, $a(\alpha) = x_\alpha^* A x_\alpha$ $(\alpha \in \mathbf{R})$. If $v^*Av > 0$, then $a(\alpha) = \alpha^2 v^*Av + 2\alpha \operatorname{Re}(u^*, Au) + u^*Au$ is positive and $b(\alpha) = \alpha^2 v^*Bv + 2\alpha \operatorname{Re}(u^*Bv) + u^*Bu$ is negative for $|\alpha|$ sufficiently large; hence $r(x_\alpha) = b(\alpha)/a(\alpha)$ is also negative for such α, and consequently $x = x_\alpha$ satisfies (3). If $v^*Av = 0$, we choose α so that $\alpha \operatorname{Re}(u^*Av) \geqslant 0$. Then $a(\alpha) = 2\alpha \operatorname{Re}(u^*Av) + u^*Au \geqslant u^*Au > 0$, and $b(\alpha)$ [and consequently also $r(x_\alpha)$] is negative if $|\alpha|$ is sufficiently large. In the remaining case, $v^*Av < 0$, it follows from hypotheses (ii) and (iii) that $0 < r(u) < r(v)$. From (i) it follows that the equation $a(\alpha) = 0$ has two real roots such that $\alpha_1 < 0 < \alpha_2$. A straightforward calculation gives, for $i = 1, 2$,

$$\lim_{\alpha \to \alpha_i} b(\alpha) = u^*Au[r(u) - r(v)] + 2\alpha_i \operatorname{Re} u^*[B - r(v)A]v.$$

Choose i such that $\alpha_i \operatorname{Re} u^*[B - r(v)A]v \leqslant 0$. From (i) we conclude that there exists $h > 0$ such that $b(\alpha) < 0$ for $\alpha \in (\alpha_i - h, \alpha_i + h)$. Noting that $a(\cdot)$ changes sign at $\alpha = \alpha_i$, it follows (after decreasing h if necessary) that $a(\alpha)$ is positive either for $\alpha \in (\alpha_i - h, \alpha_i)$ or for $\alpha \in (\alpha_i, \alpha_i + h)$. In all cases $x = x_\alpha$ satisfies (3). ■

LEMMA 3. *Let* $\mathscr{S}$ *be a subspace of dimension* $n - (d^+ + i) + 1$.

(a) *If* $\mathscr{S} \cap \{x : x^*Ax > 0\}$ *is nonempty, then there exists* $x \in \mathscr{S}$ *such that*

$$x^*Ax > 0 \quad and \quad x^*Bx \leqslant 0. \tag{4}$$

(b) *If* $\mathscr{S} \cap \{x : x^*Ax > 0\}$ *is empty, then for every* $t > 0$ *there exists* $x_t \in \mathscr{S}$ *such that*

$$x_t^*Ax_t > tx_t^*Bx_t \quad and \quad x_t^*Bx_t < 0. \tag{5}$$

The vectors x *in* (a) *and* x_t *in* (b) *can be chosen real if* A *and* B *are real and* $\mathscr{S} \cap \mathbf{R}^n$ *is a real subspace of* $\mathbf{R}^n$ *of dimension* $n - (d^+ + i) + 1$.

Proof. We assume that A and B are in their canonical forms. If $p = 0$, that is, if A is nonsingular, then we can apply Theorem 3.1 from [5]. From its proof it follows that only case (a) is possible and the existence of x satisfying (3) is given in step 2 of that proof.

For general A and B we use induction. Assume

$$A = \begin{bmatrix} A_0 & 0 \\ 0 & \tilde{A} \end{bmatrix}, \qquad B = \begin{bmatrix} B_0 & 0 \\ 0 & \tilde{B} \end{bmatrix},$$

where $A_0 = J(0, \varepsilon, m)$, $B_0 = K(\varepsilon, m)$, and $\tilde{A}$, $\tilde{B}$ are $(n-m) \times (n-m)$ matrices for which the conclusion of the lemma holds. Note that $r > 0$ in the canonical form (1) (since there exists a real eigenvalue of positive type); hence we always have nonsingular $\tilde{A}$ to start from. The pencils $\lambda A - B$ and $\lambda \tilde{A} - \tilde{B}$ have the same finite eigenvalues, and

$$\tilde{d}^+ = d^+(\tilde{A}, \tilde{B}) = d^+ - \kappa_-,$$

where $\kappa_- = \kappa_-(\varepsilon, m)$.

Let $\tilde{\mathscr{S}}$ be a subspace of $\mathbf{C}^{n-m}$ such that $\dim \tilde{\mathscr{S}} = n - m - (\tilde{d}^+ + i) + 1 = n - (d^+ + i) + 1 - \kappa_+$, where $\kappa_+ = \kappa_+(\varepsilon, m)$. According to the induction hypothesis, either there exists $x_0 \in \tilde{\mathscr{S}}$ such that $x_0^* \tilde{A} x_0 > 0$ and $x_0^* \tilde{B} x_0 \leqslant 0$, or $x^* \tilde{A} x \leqslant 0$ for all $x \in \tilde{\mathscr{S}}$ and for every $t > 0$ there exists $x_t \in \tilde{\mathscr{S}}$ such that $x_t^* B x_t < 0$ and $x_t^* \tilde{A} x_t > t x_t^* \tilde{B} x_t$.

Let $\mathscr{S}$ be a subspace of $\mathbf{C}^n$ of dimension $n - (d^+ + i) + 1$. Then

$$\mathscr{S} = \left\{x = Yz : z \in \mathbf{C}^{n-d^+-i+1}\right\},$$

where

$$Y = \begin{bmatrix} Y_0 \\ \tilde{Y} \end{bmatrix}$$

is an $n \times (n - d^+ - i + 1)$ matrix of full rank, Y_0 is an $m \times (n - d^+ - i + 1)$ matrix, and $\tilde{Y}$ is an $(n-m) \times (n - d^+ - i + 1)$ matrix.

Let $t > 0$; from Lemma 1 we find a subspace $\mathscr{T}_t$ such that $\dim \mathscr{T}_t = \kappa_-$ and that (2) holds for $0 \neq x \in \mathscr{T}_t$. Define

$$\mathscr{L}_t = \left\{z \in \mathbf{C}^{n-d^+-i+1} : Y_0 z \in \mathscr{T}_t\right\}, \qquad \tilde{\mathscr{S}} = \tilde{Y}(\mathscr{L}_t) = \left\{\tilde{Y}z : z \in \mathscr{L}_t\right\}.$$

Then $\dim \mathscr{L}_t \geqslant \dim \mathscr{T}_t + \dim \ker Y_0 \geqslant \kappa_- + n - (d^+ + i) + 1 - m = n - d^+ - i + 1 - \kappa_+$. There are two possibilities for $\mathscr{L}_t$:

1. $\ker \tilde{Y} \cap \mathscr{L}_t$ is nontrivial;
2. $\ker \tilde{Y} \cap \mathscr{L}_t = \{0\}$.

In case 1 let

$$z \in \ker \tilde{Y} \cap \mathscr{L}_t, \qquad z \neq 0, \qquad v = Yz = \begin{bmatrix} Y_0 z \\ 0 \end{bmatrix}.$$

Since Y is of full rank, $Y_0 z$ is nonzero. By definition $Y_0 z \in \mathscr{T}_t$; from (2) it follows that

$$\begin{aligned} v^*Bv &= (Y_0 z)^* B_0 (Y_0 z) < 0, \\ v^*Av &= (Y_0 z)^* A_0 (Y_0 z) > t (Y_0 z)^* B_0 (Y_0 z) = tv^*Bv. \end{aligned} \tag{6}$$

In case 2 we find $\dim \tilde{\mathscr{S}} = \dim \mathscr{L}_t \geqslant n - d^+ - i + 1 - \kappa_+$. According to the induction hypothesis one of the two alternatives holds:

(α) There exists $z_0 \in \mathscr{L}_t$ such that

$$\left(\tilde{Y}z_0\right)^* \tilde{A} \left(\tilde{Y}z_0\right) > 0 \quad \text{and} \quad \left(\tilde{Y}z_0\right)^* \tilde{B} \left(\tilde{Y}z_0\right) \leqslant 0. \tag{7}$$

(β) $(\tilde{Y}z)^* \tilde{A} (\tilde{Y}z) \leqslant 0$ for all $z \in \mathscr{L}_t$ and for every $t > 0$, there exists $z_t \in \mathscr{L}_t$ such that

$$\left(\tilde{Y}z_t\right)^* \tilde{B} \left(\tilde{Y}z_t\right) < 0 \quad \text{and} \quad \left(\tilde{Y}z_t\right)^* \tilde{A} \left(\tilde{Y}z_t\right) > t \left(\tilde{Y}z_t\right)^* \tilde{B} \left(\tilde{Y}z_t\right). \tag{8}$$

After these preparations we proceed with the induction step. We have to consider cases (a) and (b) for $\mathscr{S}$ separately.

(a): Let $u \in \mathscr{S}$ be such that $u^*Au > 0$. If $u^*Bu \leqslant 0$, then we are finished. If $u^*Bu > 0$, then fix t such that

$$0 < t < \frac{1}{r(u)} = \frac{u^*Au}{u^*Bu} \tag{9}$$

and construct the associated space $\mathscr{L}_t$.

If alternative 1 holds for $\mathscr{L}_t$, then for every $v = Yz$, $0 \neq z \in \ker \tilde{Y} \cap \mathscr{L}_t$ we find from (6) and (9) $v^*Bv < 0$ and $v^*Av > tv^*Bv > [1/r(u)]v^*Bv$, so the existence of x follows from Lemma 2.

If alternative 2 holds for $\mathscr{L}_t$, then we have to distinguish between (2α) and (2β) below:

(2α) Let $z_0 \in \mathscr{L}_t$ satisfy (7). Again there are two possibilities: if $(Yz_0)^*A(Yz_0) > 0$ then

$$(Yz_0)^*B(Yz_0) = (Y_0z_0)^*B_0(Y_0z_0) + \left(\tilde{Y}z_0\right)^*\tilde{B}\left(\tilde{Y}z_0\right)^* \leqslant (Y_0z_0)^*B_0(Y_0z_0).$$

Since $Y_0z_0 \in \mathscr{T}_t$ satisfies (2), it follows that $x = Yz_0$ satisfies (4). If $(Yz_0)^*A(Yz_0) \leqslant 0$, then (7) implies $Y_0z_0 \neq 0$ and

$$\begin{aligned} &(Yz_0)^*B(Yz_0) \leqslant (Y_0z_0)^*B_0(Y_0z_0) < 0, \\ &(Yz_0)^*B(Yz_0) < \frac{1}{t}(Y_0z_0)^*A_0(Y_0z_0) \leqslant \frac{1}{t}(Yz_0)^*A(Yz_0). \end{aligned} \tag{10}$$

From (9) and (10) we find $(Yz_0)^*B(Yz_0) < r(u)(Yz_0)^*A(Yz_0)$; therefore an application of Lemma 2 to u and $v = Yz_0$ yields x satisfying (4).

(2β) Let $z_t \in \mathscr{L}_t$ be as in (8). From $Y_0z_t \in \mathscr{L}_t$ we deduce $(Yz_t)^*B(Yz_t) < (\tilde{Y}z_t)^*\tilde{B}(\tilde{Y}z_t) < 0$ and $(Y_0z_t)^*A_0(Y_0z_t) > t(Y_0z_t)^*B_0(Y_0z_t)$; hence (8) implies

$$\begin{aligned} (Yz_t)^*A(Yz_t) &= (Y_0z_t)^*A_0(Y_0z_t) + \left(\tilde{Y}z_t\right)^*\tilde{A}(Yz_t) \\ &> t(Y_0z_t)^*B_0(Y_0z_t) + t\left(\tilde{Y}z_t\right)^*\tilde{B}(Yz_t) = t(Yz_t)^*B(Yz_t). \end{aligned}$$

Thus Yz_t satisfies

$$(Yz_t)^*B(Yz_t) < 0 \quad \text{and} \quad (Yz_t)^*A(Yz_t) > t(Yz_t)^*B(Yz_t). \tag{11}$$

Recalling (9), we can apply Lemma 2 to u and $v = Yz_t$ to find x satisfying (4).

(b): In this case $x^*Ax \leqslant 0$ for all $x \in \mathscr{S}$. Let $t > 0$. We have to find $x_t \in \mathscr{S}$ such that (5) holds.

Let $\mathscr{L}_t$ be the space associated with t as above. We consider separately cases 1 and 2 for $\mathscr{L}_t$.

In case 1 it follows from (6) that any v of the form $v = Yz$ with $0 \neq z \in \ker \tilde{Y} \cap \mathscr{L}_t$ satisfies (5).

In case 2, again we have to consider (2α) and (2β):

(2α) Since $(Yz_0)^*A(Yz_0) \leqslant 0$ (this is the property of $\mathscr{S}$), it follows from (10) that $x_t = Yz_0$ satisfies (5).

(2β) Let z_t be from (8). Then it follows from (11) that $x_t = Yz_t$ satisfies (5).

If A, B and $\mathscr{S}$ are real, then Lemmas 1 and 2 and the above proof show that x and x_t are real. ■

Proof of Theorem A. As noted earlier, we assume $\lambda_i^+ = 0$ and A, B are in their canonical form (1). Denote

$$\sigma_i = \sup_{\dim \mathscr{S} = n-(d^+ + i)+1} \inf_{\substack{x \in \mathscr{S} \\ x^*Ax>0}} r(x).$$

Let $\mathscr{S}$ be a subspace with $\dim \mathscr{S} = n - (d^+ + i) + 1$. From Lemma 3(a) it follows that $\inf\{r(x) : x \in \mathscr{S},\ x^*Ax > 0\} \leqslant \lambda_i^+ = 0$; hence

$$\sigma_i \leqslant \lambda_i^+. \tag{12}$$

Denote

$$\hat{n} = r + \sum_{j=r+1}^{r+s} m_j + 2\sum_{j=r+s+1}^{r+s+t} m_j,$$

$$\hat{d}^+ = \sum_{j=r+1}^{r+s} \kappa_+(m_j, \varepsilon_j) + \sum_{j=r+s+1}^{r+s+t} m_j,$$

and let $\hat{A}$, $\hat{B}$ be the $\hat{n} \times \hat{n}$ matrices [the "nonsingular part" of the pair (A, B)]

$$\hat{A} = \operatorname{diag}[A_1, \dots, A_{r+s+t}], \qquad \hat{B} = \operatorname{diag}[B_1, \dots, B_{r+s+t}].$$

Note that $A = \operatorname{diag}[A_{-p}, \dots, A_{-1}, \hat{A}]$, $B = \operatorname{diag}[B_{-p}, \dots, B_{-1}, \hat{B}]$, and $\hat{d}^+ = d^+(\hat{A}, \hat{B})$.

In Theorem 3.1 (step 1) of [5] we have constructed an $\hat{n} - (\hat{d}^+ + i) + 1$-dimensional subspace $\hat{\mathscr{S}}$ such that

$$\min\{r(x) : x \in \hat{\mathscr{S}},\ x^*Ax > 0\} = \lambda_i^+.$$

For $j = -1, \dots, -p$ let $\kappa_j^{\pm} = \kappa_{\pm}(m_j, \varepsilon_j)$ and define the $m_j \times m_j$ matrix

$$Y_j = \begin{array}{cc} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} & \begin{array}{c} \kappa_j^+ \\ \kappa_j^- \end{array} \\ \begin{array}{cc} \kappa_j^+ & \kappa_j^- \end{array} & \end{array}$$

Then $Y_j^* A_j Y_j = 0$ and $Y_j^* B_j Y_j \geqslant 0$. Let

$$\mathscr{S}_0 = \mathscr{R}(Y_{-p}) \oplus \cdots \oplus \mathscr{R}(Y_{-1}) \oplus \hat{\mathscr{S}}. \tag{13}$$

Then $\dim \mathscr{S}_0 = \sum_{j=-p}^{-1} \kappa_j^+ + \hat{n} - (\hat{d}^+ + i) + 1 = n - (d^+ + i) + 1$ and

$$\min\{r(x) : x \in \mathscr{S}_0,\ x^*Ax > 0\} = \min\{r(x) : x \in \hat{\mathscr{S}},\ x^*Ax > 0\} = \lambda_i^+,$$

implying $\sigma_i \geqslant \lambda_i^+$. Together with (12) this proves Theorem A. ■

In the proof of Theorem B we need another lemma.

LEMMA 4. *Let m be a positive integer, a and α real numbers, b a positive number. Then there exists a $2m \times m$ real matrix W such that $W^*K(1,2m)W$ is positive definite and $W^*[\hat{J}(a,b,m) - \alpha K(1,2m)]W$ is negative definite.*

Proof. If $m = 1$ and $w = [1 \;\; \xi]^*$, then $w^*K(1,2)w = 2\xi$, $w^*[\hat{J}(a,b,1) - \alpha K(1,2)]w = b + 2\xi(a-\alpha) - b\xi^2$. Hence w is the desired matrix if ξ is large enough.

The case $m > 1$ is reduced to $m = 1$ by perturbation. Let SR^{2m} denote the space of all $2m \times 2m$ real symmetric matrices. The set $\{H \in \mathrm{SR}^{2m} : H - \lambda K(1,2m)$ has distinct eigenvalues$\}$ is dense in SR^{2m}; this follows easily from analytic perturbation theory. The set $\{\hat{J}(a,b,m) + E : E \in \mathrm{SR}^{2m},\ 0 < E < \varepsilon I\}$ is open in SR^{2m} for $\varepsilon > 0$. Therefore we can find a positive definite E such that the pencil $\hat{J}(a,b,m) + E - \lambda K(1,2m)$ has $2m$ distinct eigenvalues $a_j \pm ib_j$, $b_j > 0$, for $j = 1,\ldots,m$. The canonical form from Section 2 implies the existence of a real nonsingular $2m \times 2m$ matrix X such that

$$X^*K(1,2m)X = \operatorname{diag}[K(1,2),\ldots,K(1,2)],$$

$$X^*\big(\hat{J}(a,b,m) + E\big)X = \operatorname{diag}\big[\hat{J}(a_1,b_1,1),\ldots,\hat{J}(a_m,b_m,1)\big].$$

From the $m = 1$ result we find $\xi_1,\ldots,\xi_m > 0$ such that $w_j = [1 \;\; \xi_j]$ satisfies $\alpha_j = w_j^* K(1,2)w_j > 0$ and $\beta_j = w_j^*[\hat{J}(a_j,b_j,1) - \alpha K(1,2)]w_j < 0$, $j =$

$1, \dots, m$. Setting

$$W_0 = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ \xi_1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & \xi_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & \xi_m \end{bmatrix}}_{m} \left.\vphantom{\begin{bmatrix} 1\\1\\1\\1\\1\\1\\1\\1 \end{bmatrix}}\right\} 2m$$

and $W = XW_0$, we obtain $W^*K(1,2m)W = \operatorname{diag}(\alpha_1, \dots, \alpha_m)$, $W^*[\hat{f}(a,b,m) + E - \alpha K(1,2m)]W = \operatorname{diag}(\beta_1, \dots, \beta_m)$. Since E is positive definite, it follows that $W^*[\hat{f}(a,b,m) - \alpha K(1,2m)]W$ is negative definite. ∎

Proof of Theorem B. Noting that the matrices Y_j in (13) and A_j, B_j, $j = -1, \dots, p$, are real, and that Lemmas 1–3 hold also in $\mathbf{R}^n$, the proof of Theorem A shows that it is sufficient to prove the theorem in the nonsingular case (i.e. $p = 0$). To do this, we repeat the two-step construction from Theorem 3.1 in [5]. The construction from step 1 can be used in the real case too, since the matrix Y on p. 226 of [5] is real and hence the real space $\mathscr{S}_i = \mathscr{R}(Y)$ has all the required properties; the matrices $\hat{Y}_j$, $j = r + s + 1, \dots, r + s + t$, can still be used, although the canonical form for A_j and B_j has been changed. The construction of the subspace $\mathscr{T}_i$ in step 2 must be slightly modified. It is constructed block by block, and there is no difference in the blocks $j = 1, \dots, r + s$, as A_j and B_j have not been changed and the matrix W_j is real if $j \leqslant r + s$. A modification is needed in the "nonreal" blocks $j = r + s + 1, \dots, r + s + t$. Instead of Lemma 3.5 in [5], which starts from the nonreal matrix J_j and yields a nonreal W_j, we apply Lemma 4, which uses the matrices from the real canonical form and yields a real matrix W_j. If we use this real W_j in the definition of W (p. 228 of [5]), then the resulting space $\mathscr{T}_i = \mathscr{R}(W)$ is real and has all the properties needed in step 2. ∎

REFERENCES

1 P. A. Binding and B. Najman, A variational principle in Krein space, *Trans. Amer. Math. Soc.*, to appear.

2 P. A. Binding and Q. Ye, Variational principles without definiteness conditions, *SIAM J. Math. Anal.* 22:1575–1583 (1991).
3 I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products*, Birkhäuser, Basel, 1983.
4 P. Lancaster and Q. Ye, Variational properties and Rayleigh quotient algorithms for symmetric matrix pencils, in *Oper. Theory Adv. Appl.* 40, Birkhäuser, Basel, 1989, pp. 247–278.
5 B. Najman and Q. Ye, A minimax characterization for eigenvalues of hermitian pencils, *Linear Algebra Appl.* 144:217–230 (1991).
6 R. C. Thompson, Pencils of complex and real symmetric and skew matrices, *Linear Algebra Appl.* 147:323–371 (1991).
7 F. Uhlig, A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil, *Linear Algebra Appl.* 14:189–209 (1976).

Received 7 July 1992; final manuscript accepted 13 October 1992